A new nonlocal thermodynamical equilibrium radiative transfer method for cool stars

Method & numerical implementation

J. Lambert[1,2], E. Josselin[2], N. Ryde[1], and A. Faure[3]

[1] Lund Observatory, Box 43, SE-221 00 Lund, SWEDEN
[2] Laboratoire Univers et Particules de Montpellier (LUPM), UMR 5299, Université Montpellier 2 , CNRS
  Université Montpellier 2 - CC72 Place Eugène Bataillon 34095 Montpellier Cedex 5 FRANCE
[3] Institut de Planètologie et d'Astrophysique de Grenoble (IPAG), UMR 5109, Université Joseph Fourier, CNRS, OSUG,
  38041 Grenoble Cedex 9, FRANCE
  e-mail: julien.lambert@astro.lu.se

ABSTRACT

Context. The solution of the nonlocal thermodynamical equilibrium (non-LTE) radiative transfer equation usually relies on stationary iterative methods, which may falsely converge in some cases. Furthermore, these methods are often unable to handle large-scale systems, such as molecular spectra emerging from, for example, cool stellar atmospheres.
Aims. Our objective is to develop a new method, which aims to circumvent these problems, using nonstationary numerical techniques and taking advantage of parallel computers.
Methods. The technique we develop may be seen as a generalization of the coupled escape probability method. It solves the statistical equilibrium equations in all layers of a discretized model simultaneously. The numerical scheme adopted is based on the generalized minimum residual method.
Results. The code has already been applied to the special case of the water spectrum in a red supergiant stellar atmosphere. This demonstrates the fast convergence of this method, and opens the way to a wide variety of astrophysical problems.

Key words. Radiative transfer – Methods: numerical – Stars: atmospheres

1. Introduction

Radiative transfer plays a central role in astrophysics. The solution of the radiative transfer equation (hereafter RTE) is required to properly interpret any spectrum observed from an astrophysical object. Furthermore the radiation field often represents an important ingredient in hydrodynamics, for instance, through the energy budget (radiative heating and/or cooling) and the radiative pressure. However, as we discuss below, while the standard expression of the RTE is an ordinary differential equation of the 1st order, it may be highly nonlinear and nonlocal, especially because of scattering processes. The solution of the RTE thus requires specific procedures. Furthermore, with the development of infrared (IR) astronomy (e.g., from CRIRES@VLT in the near-IR to ALMA through HIFI/Herschel), one is now often confronted with very large-scale problems resulting from molecular spectra, which most techniques developed so far cannot handle.

The techniques for solving the RTE may be divided into a few strategies. Among approximate methods, one that is widely used, relies on the escape probability technique (e.g., Castor 1970), which does not take the nonlocality of the RTE into account. Secondly, Monte Carlo methods can be used because of the probabilistic nature of scattering processes. These methods can successfully be adapted to any geometry, but may be unadapted when large optical depths are met (see, e.g., Juvela 2005, for a recent review). We do not discuss these two classes of methods hereafter.

To properly treat the nonlocality and the nonlinearity of the RTE, it is usually rewritten in the form of a diffusion integral equation, which is then solved iteratively[1]. This integral form is usually referred to as the Λ operator. The form involves a kernel, which determines the mathematical nature of the integral equation. For example, in the case of an assumed plane parallel geometry, the kernel involves the exponential integral function, and the RTE then becomes a weakly singular Fredholm equation of the second type (this singularity is an important aspect, which will be treated below). It can then be formally solved thanks to appropriate algorithms, such as the Atkinson algorithm (Ahues et al. 2002).Since the inversion of the Λ operator is usually numerically prohibitive, one needs an iteration scheme. It is well known that the iteration with the complete Λ operator at best converges very slowly or even stabilizes far from the real solution. A major step in solving the RTE has been the introduction of the so-called approximate lambda iteration (ALI), initially by Cannon (1973), then called the operator splitting method. The ALI is in fact an application of the Jacobi method to the RTE, as shown by Olson et al. (1986), who greatly improved the mathematical understanding of the problem. In these methods, the choice of the approximate operator (denoted $\Lambda^{\star}$) is crucial. The

---

[1] We already emphasize here that solving either the RTE explicitly or the statistical equilibrium equation is strictly identical, and both equations are of similar mathematical nature, as shown, for instance, in the Appendix of Hauschildt et al. (1995)

most common choice for $\Lambda^\star$ is the diagonal part of Λ, which ensures a rapid convergence if Λ is diagonally dominant matrix (Gershgorin's Theorem). ALI has been applied to a large variety of problems, including 3D polarized transfer (Štěpán & Trujillo Bueno 2013). This latter work also emphasizes the potential progress that can be made by taking advantage of massively parallel computing.

Meanwhile, new mathematical techniques have been applied to the RTE problem, especially the Gauss-Seidel method and the successive over relaxation (SOR) method, which significantly improve the convergence properties (Trujillo Bueno & Fabiani Bendicho 1995) that are determined by the spectral radius of their amplification matrix. Besides these, complete linearization methods have also been developed, the well-known case being MULTI (Scharmer & Carlsson 1985). Despite apparent different formulations, the two approaches are essentially the same, as shown by Socas-Navarro & Trujillo Bueno (1997).

Indeed, all these methods are based on linearization techniques. In particular, the dependence of the radiation field on the populations of the energy levels is approximated. Furthermore, all these methods are stationary (in the case of SOR, this is true if the relaxation parameter is constant over iterations), meaning that the spectral radius is constant. As the error vector $\mathbf{e}^m$ at the mth iteration is proportional to the power m of the spectral radius of the amplification matrix $|\lambda_{\max}|^m$, the convergence may thus be very slow in some cases. For the Lambda iteration, the spectral radius is close to 1 when the photon destruction probability $\epsilon \ll 1$ and/or the total optical depth is $\gg 1$, explaining why it falsely converges. This is illustrated in Figs. 1 and 2, where the classical Eddington problem (two-level atom with constant ϵ) is solved with Λ iteration (LI) and ALI, respectively. The false convergence of LI and the very low convergence of ALI are illustrated in cases B. Indeed, their rate of convergence is very slow for strongly nonlocal thermodynamical equilibrium (non-LTE) problems (non-LTE parameter $\epsilon \ll 1$, ϵ being the fraction of the source function given by the Planck function, or the photon destruction probability) and to the number of points per decade of optical depth. In the latter case (i.e., a fine sampling of the stellar atmosphere), this problem of slow convergence is well explained by Fabiani Bendicho et al. (1997), who also showed that the multigrid method can nicely circumvent this problem. This fine sampling is however rarely required, especially as classical model atmospheres are static. An excellent discussion of the general problem of false convergence is given in Chapter 13 in Hubeny & Mihalas (2014).

The iterative schemes are usually combined with a technique of acceleration of convergence (Auer 1991). These fall into two main categories. The first is based on the minimization of the residual, which is the case of the Ng acceleration (Ng 1974). The second class relies on minimization with respect to a set of conjugate vectors. This was first applied to the RTE by (Klein et al. 1989 ; see also Dickel & Auer 1994).

The second class of techniques (the minimization with respect to a set of conjugate vectors) leads naturally to more general, nonlinear, nonstationary methods, such as the conjugate gradient and the generalized minimum residual method (GMRES). To our knowledge, only the conjugate gradient has been applied to the RTE, by Paletou & Anterrieu (2009).

The aim of our work is to develop a new nonstationary method for solving the RTE, which fully takes its nonlinearity into account. This new method thus opens the possibility to explore the mathematical consequences of this intrinsic property of the RTE. Furthermore, circumventing the possible slow or false convergence of stationary methods should allow an appro-
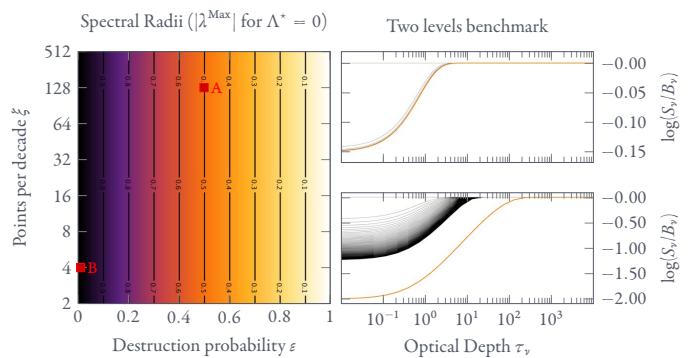
Fig. 1. Spectral radius of amplification matrix for the Λ iteration method. Convergence is shown for two spectral radii, marked with letters A (upper panel) and B (lower panel) in the right-hand panels. The thin black lines correspond to successive iterations; the true solution is given by the orange line.
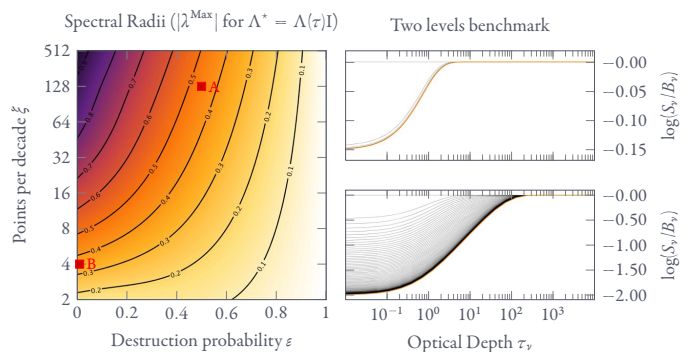

Fig. 2. Same as Figure 1, except for ALI with $\Lambda^\star$ = diagonal of Λ.

priate treatment of some specific cases, such as, e.g., shocked regions, which require a fine sampling and present a large dynamics of optical depth values. Finally, as the coding strategy of this method has been conceived as a parallel code from the beginning, it is applicable to large-scale systems such as molecular spectra, and this was indeed our initial motivation (a posteriori parallelization is often less efficient, or even impossible).

The derivation of the equations, including the Jacobian matrix coefficients, used in the nonlinear method is presented in section 2. The implementation of this method and its parallelization are described in section 3. The application of the code to a specific problem facilitates the determination of its convergence properties, as shown in section 4.

2. Global strategy

Contrary to the two-level atom problem with a constant destruction probability, which can be directly solved by direct inversion (or through the analytic Eddington solution). Multi-level atom or molecule problems require an iterative determination for two reasons. Firstly, the multi-level problem is nonlinear. Secondly, the radiative transfer equation is nonlocal, and one has to propagate the modification of the field during iterations. As we discussed in the previous section, however, this stationary iteration may be problematic for large-scale systems. The basic concept of the strategy we adopt is to suppress the iterations due to the nonlocality of the RTE by considering the full system (spatially) and computing the radiation field exactly (assuming given populations) with an analytic formulation of the mean intensity field, which explicitly depends on populations.

2.1. Formulation of the problem

For clarity, we hereafter consider the special case of plane parallel geometry. In this context the 1D RTE is

$$\mu \frac{dI_\nu}{ds} = \eta_\nu^{\text{Cont}} + \eta_\nu^{\text{Line}} - (\chi_\nu^{\text{Cont}} + \chi_\nu^{\text{Line}})I_\nu, \tag{1}$$

where I_ν is the specific intensity, η and χ are the emission and extinction coefficients, and s is the geometrical path. We explicitly separate the continuum and the line processes (all demonstrations and symbols can be found in Appendix A and C, respectively). Assuming complete redistribution, we introduce classical expressions for line processes,

$$\chi_\nu^{\text{Line}} = \phi_\nu^{ul}\chi_{ul} = \phi_\nu^{ul}\frac{h\nu_{ul}}{4\pi}(B_{lu}n_l - B_{ul}n_u) \tag{2a}$$

$$\eta_\nu^{\text{Line}} = \phi_\nu^{ul}\eta_{ul} = \phi_\nu^{ul}\frac{h\nu_{ul}}{4\pi}A_{ul}n_u, \tag{2b}$$

where u and l refer to the upper and lower levels, respectively. Hereafter, if the distinction between upper and lower levels is not required, the indices i and j are used. The normalized line profile is ϕ_ν^{ul} and A and B are the Einstein coefficients.

Following the strategy of Gonzalez Garcia et al. (2008), we then introduce a unique optical depth scale for all frequencies, e.g,. the commonly used τ_V scale (optical depth at 500 nm). Eq. (1) then becomes

$$\mu \frac{dI_\nu}{d\tau_V} = \left[\xi_\nu + E_{ij}\zeta\phi_\nu^{ij}(x_j - x_i)\right]I_\nu - D_{ij}x_i\zeta\phi_\nu^{ij} - \xi_\nu S_\nu, \tag{3}$$

where

$$\xi_\nu = \frac{\tilde{\chi}_{\nu_{ij}}^{\text{Cont}}}{\tilde{\chi}_V} \tag{4a}$$

$$\zeta = \frac{1}{\tilde{\chi}_V}\frac{n^k}{n^H} \tag{4b}$$

$$E_{ij} = \frac{h\nu_{ij}}{4\pi}B_{ij}g_i \tag{4c}$$

$$D_{ij} = \frac{h\nu_{ij}}{4\pi}A_{ij}g_i \tag{4d}$$

$$x_i = \frac{n_i}{n^k}\frac{1}{g_i}. \tag{4e}$$

The notation $\tilde{}$ means that the quantity is expressed per hydrogen atom. Details can be found in Appendix A. As implicitly assumed in Eq. 3, we do not consider line overlap. This assumption is justified in the case of infrared molecular lines. Otherwise, a summation over transitions (at least adjacent ones) would be required.

This formulation depends explicitly on the populations of each energy level of the species considered. These populations are solutions of the statistical equilibrium equation system [2], which is for each level n_i, the detailed balance of incoming and outgoing (de-)excitation processes, defined by radiative (Einstein coefficients), and collisional rates (n_{col} being the local density of the colliding partner). Thus the statistical equilibrium equations take the form

$$\frac{dn_i}{dt} = n_i\left[\sum_{j<i}A_{ij} + \sum_{j\neq i}\left(B_{ij}\bar{\mathcal{J}}_{ij} + n_{col}C_{ij}\right)\right]$$

$$- \left[\sum_{j>i}n_jA_{ji} + \sum_{j\neq i}n_j\left(B_{ji}\bar{\mathcal{J}}_{ji} + n_{col}C_{ji}\right)\right] = 0, \tag{5}$$

---

[2] At steady-state. Chemical formation and destruction terms are not included.

where the mean radiation field $\bar{\mathcal{J}}_{ij}(\tau_V) = \bar{\mathcal{J}}_{ji}(\tau_V)$ depends on the solution of the RTE,

$$\bar{\mathcal{J}}_{ij}(\tau_V) = \frac{1}{2}\int_0^\infty \phi_\nu^{ij}(\tau_V)\int_{-1}^{+1}I_\nu(\tau_V, \mu)d\mu d\nu. \tag{6}$$

Because this system of equations is degenerate, one of them is replaced by the normalization equation $\sum_i^N n_i = n^k$, where n^k is the total density of the species k.

To express each equation of the statistical equilibrium system with the same unknowns as in Eq. 3, we rewrite the statistical equilibrium with the variables x_i defined in Eq. 4. Moreover, for numerical reasons, we make a variable change for the mean intensity,

$$\bar{\mathcal{J}}_{ij}(\tau_V) = \frac{2h\nu_{ij}^3}{c^2}\bar{\mathcal{Z}}_{ij}(\tau_V). \tag{7}$$

This leads to a matrix formulation of the statistical equilibrium, with an upper triangular part corresponding to incoming processes, and a lower triangular part corresponding to outgoing processes. Finally, thanks to the Einstein relations ($B_{ul} = (c^2/(2h\nu_{ul}^3))A_{ul}$ and $B_{lu} = (g_u/g_l)B_{ul}$), the system of equations (5) becomes

$$x_i\left[\sum_{j<i}g_iA_{ij}\left(1 + \bar{\mathcal{Z}}_{ij}\right) + \sum_{j>i}g_jA_{ji}\bar{\mathcal{Z}}_{ij} + \sum_{j\neq i}n_{col}g_iC_{ij}\right]$$

$$- \left[\sum_{j<i}x_jg_jA_{ij}\bar{\mathcal{Z}}_{ij} + \sum_{j>i}x_jg_jA_{ji}\left(1 + \bar{\mathcal{Z}}_{ji}\right) + \sum_{j\neq i}n_{col}g_iC_{ij}\right] = 0. \tag{8}$$

Because radiative transfer is a nonlocal problem, an integral formulation of the mean radiation field with an explicit dependence on populations is required.

2.2. The mean radiation field

To derive the mean radiation field, the first step consists of deriving the formal solution of the RTE. A classical way to obtain an integral form of the formal solution of a differential equation is to use the Green's functions $\mathcal{G}_{\mu\nu}$. This method consists of replacing the source term by a Dirac function. Taking the boundary conditions into account, the equation is then solved using its Laplace transform. The complete solution is then obtained after integration of this Green's function, that is, performing a summation over each point source, according to the superposition principle (which is valid if the differential operator is linear). Thus, the integral form of $I_\nu(\tau_V, \mu)$ can be obtained by the determination of the Green's function of Eq. (3). This Green's function is easily computable for spherical and other geometries using some symbolic tools such as Mathematica, making it possible to generalize the method for 2D or 3D problems. Knowing this Green's function, we can directly obtain the formal integral form of the specific intensity, which is

$$I_\nu(\tau_V, \mu) = \int_{\tau_V^{\text{Min}}}^{\tau_V^{\text{Max}}}\left[D_{ij}x_i(t)\zeta(t)\phi_\nu^{ij}(t) + \xi_\nu(t)S_\nu(t)\right]\mathcal{G}_{\mu\nu}(\tau_V; t)dt. \tag{9}$$

We assume the line profile does not depend on μ. This expression is thus not applicable in the case of anisotropic scattering or in the presence of radiation fields, but is valid in the case of

classical static model atmospheres. The formal expression of the mean intensity, $\bar{\mathcal{J}}_{ij}(\tau_V)$, requires an integration over solid angles and frequencies according to Eq (6). We assume that $D_{ij}x_i\zeta$ and $\xi_\nu S_\nu$ are angular independent (physically this corresponds to the assumption of isotropic scattering) and frequency independent over the line width ($\xi_\nu S_\nu \rightarrow \xi_{ij}S_{\nu_{ij}}$). This is a very weak assumption because the continuum is nearly constant on the scale of a line width. Then, we can extract the source terms from the integral over solid angles and frequencies. Moreover, we separate line and continuum contributions by splitting the integral, which leads to

$$\bar{\mathcal{J}}_{ij}(\tau_V) = \int_{\tau_V^{Min}}^{\tau_V^{Max}} \xi_{ij}(t)S_{\nu_{ij}}(t)$$
$$\times \left[\frac{1}{2}\int_0^\infty \phi_\nu^{ij}(\tau_V)\left[\int_{-1}^{+1}\mathcal{G}_{\mu\nu}(\tau_V;t)d\mu\right]d\nu\right]dt$$
$$+ D_{ij}\int_{\tau_V^{Min}}^{\tau_V^{Max}} x_i(t)\zeta(t)$$
$$\times \left[\frac{1}{2}\int_0^\infty \phi_\nu^{ij}(\tau_V)\phi_\nu^{ij}(t)\left[\int_{-1}^{+1}\mathcal{G}_{\mu\nu}(\tau_V;t)d\mu\right]d\nu\right]dt. \quad (10)$$

Then, after determination of the Green's function in a given geometry, one obtains an integral form for the different contributors to the radiation field. We assume that the continuum contribution is in LTE ($S_{\nu_{ij}} = B_{\nu_{ij}}(T)$). Moreover, we consider, as boundary conditions, an incoming field I_-^{Ext} from the inner atmosphere. Then,

$$\bar{\mathcal{J}}_{ij}(\tau_V) = \bar{\mathcal{J}}_{ij}^{Ext}(\tau_V) + \bar{\mathcal{J}}_{ij}^{Cont}(\tau_V) + \bar{\mathcal{J}}_{ij}^{Line}(\tau_V) \quad (11)$$

where,

$$\bar{\mathcal{J}}_{ij}^{Ext}[x_i, x_j](\tau_V) = I_-^{Ext}\beta_-[x_i, x_j](\tau_V) \quad (12a)$$

$$\bar{\mathcal{J}}_{ij}^{Cont}[x_i, x_j](\tau_V) = \int_{\tau_V^{Min}}^{\tau_V^{Max}} \xi_{ij}(t)B_{\nu_{ij}}(T(t))L_1[x_i, x_j](t, \tau_V)dt \quad (12b)$$

$$\bar{\mathcal{J}}_{ij}^{Line}[x_i, x_j](\tau_V) = D_{ij}\int_{\tau_V^{Min}}^{\tau_V^{Max}} x_i(t)\zeta(t)K_1[x_i, x_j](t, \tau_V)dt, \quad (12c)$$

where $\beta_-[x_i, x_j](\tau_V)$ is the escape probability from position τ_V in the atmosphere,

$$\beta_-[x_i, x_j](\tau_V) = \frac{1}{2}\int_0^{+\infty} \phi_\nu^{ij}(\tau_V) \times E_2\left(\tau^{Tot}[x_i, x_j](\tau_V)\right)d\nu. \quad (13)$$

The kernels $L_1[x_i, x_j](t, \tau_V)$ and $K_1[x_i, x_j](t, \tau_V)$ are, respectively:

$$L_1[x_i, x_j](t, \tau_V) = \frac{1}{2}\int_0^{+\infty} \phi_\nu^{ij}(\tau_V)$$
$$\times E_1\left(-\varepsilon(t, \tau_V)(\tau^{Tot}[x_i, x_j](t) - \tau^{Tot}[x_i, x_j](\tau_V))\right)d\nu \quad (14a)$$

$$K_1[x_i, x_j](t, \tau_V) = \frac{1}{2}\int_0^{+\infty} \phi_\nu^{ij}(\tau_V)\phi_\nu^{ij}(t)$$
$$\times E_1\left(-\varepsilon(t, \tau_V)(\tau^{Tot}[x_i, x_j](t) - \tau^{Tot}[x_i, x_j](\tau_V))\right)d\nu, \quad (14b)$$

where the total optical depth is given by

$$\tau_\nu^{Tot}(\tau_V) = \int_0^{\tau_V}\left(\chi_\nu^{Cont} + \chi_\nu^{Line}\right)dt$$
$$= \int_0^{\tau_V}\left(\xi_\nu(t) + E_{ij}\zeta(t)\phi_\nu^{ij}(t)[x_j(t) - x_i(t)]\right)dt. \quad (15)$$

We do not include an external radiation field. In the plane parallel geometry considered here, it would require us to take additional effects such as limb darkening in realistic cases into account (e.g., illumination by a planet or a companion). We have introduced the function $\varepsilon(t, \tau_V) = 1$ if $t < \tau_V$ and -1 if $t > \tau_V$ to split the integration over μ according to its sign. The physical meaning of the kernels $L_1(t, \tau_V)$ and $K_1(t, \tau_V)$ can be understood as the probability density for photons, emitted in a layer t, to contribute to the line excitation in the layer τ_V.

We now have an analytic expression of the mean intensity with an explicit dependence on populations. We now have to include this expression in the equations of statistical equilibrium to close the system and form a system of coupled integral equations (a Fredholm-like equation system).

2.3. The multi-zone statistical equilibrium equation system

The inversion of the system of statistical equilibrium equations allows us to obtain the populations also including the formal solution of RTE. But because it is nonlocal, the knowledge of $\bar{\mathcal{J}}$ requires the knowledge of the populations everywhere in the atmosphere because of the integral over the optical depth t in Eq. (12).

For this reason, one needs to solve the statistical equilibrium everywhere in the atmosphere simultaneously. The first step consists of rewriting Eq. (5) in a matrix form. A vector $\mathbf{x}(\tau_V)$ containing all the populations of the considered species in a given layer, normalized to the density in that layer, is built. The size of this vector is thus equal to N, the number of energy levels of the considered species. We introduce the rate matrix

$$\mathbf{R}[\mathbf{x}](\tau_V) = \mathbf{A}^T \circ \left(\mathbf{L}_1 + \bar{\mathbf{Z}}[\mathbf{x}](\tau_V)\right)^T$$
$$+ \text{diag}(\mathbf{g}) \cdot \left[\mathbf{A} \circ \bar{\mathbf{Z}}[\mathbf{x}](\tau_V)\right] \cdot \text{diag}(\mathbf{g})^{-1} + \mathbf{C}^T(\tau_V). \quad (16)$$

Where $\mathbf{L}_1$ is a lower triangular matrix of ones, $\mathbf{A}$ is a vector containing the spontaneous de-excitation coefficients (A_{ul}), $\mathbf{Z}$ is a matrix containing the radiation field (see Eq. 7), $\mathbf{C}$ a matrix containing the collisional rates ($n_{col}C_{ij}$), and $\mathbf{g}$ is a vector containing the statistical weights (g_i).

Now the statistical equilibrium system can be written as,

$$\mathbf{M}^T[\mathbf{x}](\tau_V)\mathbf{x}(\tau_V) = \mathbf{0}, \quad (17)$$

with

$$\mathbf{M}[\mathbf{x}](\tau_V) = [\text{diag}(\mathbf{u} \cdot \mathbf{R}[\mathbf{x}](\tau_V)) - \mathbf{R}[\mathbf{x}](\tau_V)] \cdot \text{diag}(\mathbf{g}), \quad (18)$$

where $\mathbf{u}$ is a ones vector. After discretization of the medium into P layers indexed by l, the vectorial function $\mathbf{x}(\tau_V)$ becomes $\mathbf{x}^l$ and equation (16) becomes

$$\mathbf{R}^l(\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^P) = \mathbf{A}^T \circ \left(\mathbf{L}_1 + \bar{\mathbf{Z}}^l(\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^P)\right)^T$$
$$+ \text{diag}(\mathbf{g}) \cdot \left[\mathbf{A} \circ \bar{\mathbf{Z}}^l(\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^P)\right] \cdot \text{diag}(\mathbf{g})^{-1} + \left[\mathbf{C}^l\right]^T \quad (19)$$

as the radiation field, and thus the matrix $\mathbf{Z}$ depends implicitly on the populations in all layers , and the statistical equilibrium system can be rewritten,

$$\left[\mathbf{M}^l(\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^P)\right]^T \mathbf{x}^l = \mathbf{0}. \quad (20)$$

To take care of the coupling between the layers and close the system to solve Eq. (20) for all l in a unique step, we build the

multi-zone statistical equilibrium system:

$$
\begin{pmatrix}
\mathbf{M}^1(\mathbf{X}) & 0 & 0 \\
0 & \ddots & 0 \\
0 & 0 & \mathbf{M}^P(\mathbf{X})
\end{pmatrix}^T
\begin{pmatrix}
\mathbf{x}^1 \\
\vdots \\
\mathbf{x}^P
\end{pmatrix}
=
\begin{pmatrix}
\mathbf{0} \\
\vdots \\
\mathbf{0}
\end{pmatrix}.
\tag{21}
$$

We introduce the vector $\mathbf{X} = (\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^P)$ and the block-diagonal matrix $\mathbf{\Gamma}$ to rewrite Eq. (21) as a multi-D nonlinear function,

$$
[\mathbf{\Gamma}(\mathbf{X})]^T \mathbf{X} = \mathbf{F}(\mathbf{X}) = \mathbf{0}.
\tag{22}
$$

With an initial $\mathbf{X}_0$ within the convergence radius, we can solve this nonlinear equation with a classical scheme such as the Newton method. An important question, however, is to determine whether the above system has a solution and whether the solution we expect to get is physical (unique and positive).

2.4. Existence, uniqueness and positivity of the solution

From a physical point of view, one naturally expects a unique and positive solution of the RTE. Reflecting on this problem is in fact a way to verify whether it is mathematically well formulated. Indeed, as stated above, the coupling between the populations and the radiation field is nonlinear. Without linearization, the solution can be regarded as the equilibrium state of Markov processes. Furthermore, because of radiative selection rules and collisional propensity rules, many factors are very small, and may numerically tend toward zero. The problem is thus potentially ill-conditioned.

The formulation adopted here is a multi-zone statistical equilibrium. The problem of uniqueness and positivity of the solution of the mono-zone statistical equilibrium has been addressed by Damgaard et al. (1992) and Rybicki (1997). As this system of equations is degenerate, one equation is usually replaced by a normalization condition, namely the conservation equation, to overcome the singularity of the associated matrix. Then, Damgaard et al. (1992) show that the solution is positive if all the rates are strictly non-zero. Yet certain transition probabilities vanish, physically or numerically. As long as all the levels are connected, just one nonzero cofactor guarantees the regularity of the matrix. More generally, Rybicki (1997) demonstrated that even for the system without a normalization condition, positive states can link (and be linked) only to positive states. Then, if the statistical equilibrium equations can be transformed into a set of uncoupled irreducible subproblems with a strictly positive normalization condition, the solution is positive and unique. In other words, the uniqueness and positivity properties of the general *linear* statistical equilibrium equations relate to the underlying connectivity properties of the states.

In our multi-zone formulation, the coupling between statistical equilibria in every atmospheric layer is strongly nonlinear. That is, the system cannot be split into uncoupled sets of equations, i.e., a set of uncoupled irreducible subproblems. This may impact the number of required normalization conditions, while only one per atmospheric layer (the conservation equation) is available. The uniqueness of the solution may then be demonstrated through the search for a nonlinear fixed point as the solution of the equilibrium of a Markov chain (P. Azerad, priv. comm.). A forthcoming paper will be devoted to this. At least, as the properties of connectivity still apply, the positivity of the solution is guaranteed.

3. Implementation

3.1. Nonlinear solving

In the method developed here, the computation of the statistical equilibrium is based on its exact formulation and explicitly includes radiative transfer effects through the mean radiation field $\bar{\mathcal{J}}$ in equation (22). Finding the root of this function naturally leads to the exact solution to the statistical equilibrium, and avoids any false convergence that may be obtained with stationary iterative methods.

The computation of the norm of a function,

$$
\mathbf{F} : (\mathbb{R}^+)^{N \times P} \to \mathbb{R}^{N \times P},
\tag{23}
$$

N and P being the number of considered energy levels and the number of layers in the model atmosphere, respectively, is equivalent to an optimization problem, that is, the search for the minimum of the L_2 norm of the function f,

$$
\min_{\mathbf{X} \in \mathbb{R}^{N \times P}} f(\mathbf{X}) = \frac{1}{2}\mathbf{F}(\mathbf{X})\mathbf{F}^T(\mathbf{X}) = \frac{1}{2}\|\mathbf{F}(\mathbf{X})\|^2_{L_2}.
\tag{24}
$$

The classical way to minimize this function is to move along a vector $\delta\mathbf{X}$ in a descending direction, using the Newton's method. Indeed, for $\delta\mathbf{X} = \mathbf{X}_k - \mathbf{X}_{k-1} = -\mathbf{J}^{-1} \cdot \mathbf{F}$, the gradient along the vector $\delta\mathbf{X}$ is

$$
\begin{aligned}
\nabla f(\mathbf{X}_k)^T \delta\mathbf{X} &= \left[\mathbf{F}(\mathbf{X}_k)^T \mathbf{J}(\mathbf{X}_k)\right]\left[-\mathbf{J}(\mathbf{X}_k)^{-1}\mathbf{F}(\mathbf{X}_k)\right] \\
&= -\mathbf{F}(\mathbf{X}_k)^T \mathbf{F}(\mathbf{X}_k) = -\|\mathbf{F}(\mathbf{X})\|^2_{L_2} < 0.
\end{aligned}
\tag{25}
$$

Newton's method is very efficient in approaching the solution as it has a quadratic convergence. For each successive Newton step, we have to solve the linear problem $\mathbf{J}\delta\mathbf{X}_k = -\mathbf{F}$, where k is the index of the Newton iteration. However, with classical algebric methods, this operation leads to an $O(N^3)$ algorithm, which is prohibitive in terms of computational time for large-scale systems. A robust and efficient way to accelerate this operation is to approximate the solution using a faster algorithm. We adopt GMRES, which is a subclass of Krylov subspace methods.

A Krylov subspace method begins with an initial guess $\delta\mathbf{X}_k^0$. At the nth sub-iteration, $\delta\mathbf{X}_k^n$ is determined through a correction in the nth Krylov subspace,

$$
\mathcal{K}_k^n = \text{span}(\mathbf{R}, \mathbf{J}\mathbf{R}, ..., \mathbf{J}^{n-1}\mathbf{R}); \ \mathbf{R} = -\mathbf{F} - \mathbf{J}\delta\mathbf{X}_k^n.
\tag{26}
$$

The main advantage of this method is that its implementation requires only some vector matrix (or vector transposed matrix) products. It leads to a complexity in $O(n^2)$, which may be reduced to $O(n)$ if the matrix is sparse. Moreover, the scheme needs $O(kn)$ operations, but the algorithm can be restarted to minimize this contribution.

For each Newton step, we thus have to solve the linear problem $\mathbf{J}\delta\mathbf{X} = -\mathbf{F}$, which will be rewritten in this section as $\mathbf{A}\mathbf{x} = \mathbf{b}$ for simplicity. Stationary iterative methods (Jacobi, Gauss-Seidel, SOR) approximate the unknown vector $\mathbf{x}$ with a scheme $\mathbf{x}^{(k)} = B\mathbf{x}^{(k-1)} + \mathbf{c}$, where neither B nor $\mathbf{c}$ depend upon the iteration count k. We use the GMRES method, which is part of a class of (nonstationary) methods named Krylov subspace methods. The GMRES algorithm tries to evaluate the guess $\mathbf{x}_n \in \mathcal{K}_n$ (the n-th order Krylov subspace) by solving an optimization problem (minimizing the residue $\|\mathbf{A}\mathbf{x}_n - \mathbf{b}\|_{L_2}$) like a least-square problem. The n-th Krylov subspace is spanned over the n basis vectors,

$$
\mathcal{K}_n = \text{span}(\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}..., \mathbf{A}^{n-1}\mathbf{b}).
\tag{27}
$$

Then, the approximate solution is $\mathbf{x}_n = \mathbf{K}_n\mathbf{c}$, where the Krylov matrix is $\mathbf{K}_n = \begin{bmatrix} \mathbf{b}, \mathbf{Ab}, \mathbf{A}^2\mathbf{b}..., \mathbf{A}^{n-1}\mathbf{b} \end{bmatrix}$ and $\mathbf{c}$ is an appropriate vector such that

$$\min \|\mathbf{Ax}_n - \mathbf{b}\|_{L_2} = \min \|\mathbf{AK}_n\mathbf{c} - \mathbf{b}\|_{L_2}. \tag{28}$$

This vector is easily computed as it just requires a Matrix-vector product. This is particularly well suited for sparse matrices because this operation evolves roughly as the number of nonzero elements.

A straightforward method to find the least-square solution of this problem would be to compute the QR factorization of the matrix $\mathbf{AK}_n$. Indeed, the normal equation $A^T A x = A^T b \Leftrightarrow R x = Q^T b$ can be solved by back substitution. However, this is both unstable and too expensive (the QR factorization requires $2mn^2$ flops, $d = Q^T b$, $2mn$ flops, and $Rx = d$, n^2 flops, where m is the iterate and n is the order of the Krylov subspace). Moreover this natural Krylov basis is ill-conditioned and is thus not a good choice. Indeed, the basis vectors have to be linearly independent, and the successive multiplications by A lead to vectors, which point (fastly) to the dominant eigenvector of A and are thus numerically collinear. For this reason, we build another space , Q_n , where the basis vectors are orthogonal (this orthogonalization of Krylov bases is named the Arnoldi method). During the Arnoldi step to build the orthogonal basis, the matrix H of the orthogonalization elements is built. This matrix has a special shape (Hessenberg shape), which is close to triangular (one subdiagonal more). Our minimization problem can be rewritten as

$$\min \|\mathbf{Ax}_n - \mathbf{b}\|_{L_2} = \min \|\mathbf{AQ}_n\mathbf{y} - \mathbf{b}\|_{L_2} = \min \|\mathbf{Q}_{n+1}\mathbf{H}_{n+1}\mathbf{y} - \mathbf{b}\|_{L_2}, \tag{29}$$

where $\mathbf{y}$ designs the new unknown ($\mathbf{Q}_n\mathbf{y} = \mathbf{x}_n$).

Furthermore, we have to combine this method with a global convergence strategy. Indeed, as one gets close to the solution, the choice of the descending direction may lead to spurious effects, especially if $\delta\mathbf{X}$ becomes very large. Then, to ensure the global convergence of the scheme, we use a combination of the line search and backtracking methods, which is an efficient method to solve nonlinear equations. Precisely, when $f_k > f_{k-1}$, $\delta\mathbf{X}$ may be too large. Then, $\delta\mathbf{X}$ is reduced by a factor λ,

$$\mathbf{X}_k = \mathbf{X}_{k-1} + \lambda\delta\mathbf{X}, \tag{30}$$

with λ chosen to minimize $f(\mathbf{X}_{k-1} + \lambda\delta\mathbf{X})$ to maintain the iteration in the descending direction.

3.2. Treatment of the singularity

The computation of $\bar{\mathcal{J}}$ goes through the integration of a singular function. Indeed, in the integrand of Eq. (12), the function $E_1(x)$ has a logarithmic divergence at $x = 0^+$, see Fig 3.

To integrate this kernel, we use the periodization method developed by Helluy et al. (1998), which consists of a high order quadrature method with a variable change. The new variable is a polynomial of degree k, properties of which improve on the order of rectangle rule. The advantage of this method is that it can handle singular functions. Indeed the error $|E_N(f)|$ on an interval subdivided in N subintervals goes as $\sim C_k/N^\gamma$, where $\gamma \to (2k - 1)$ if f has a logarithmic singularity at one bound of the integration domain (see Table 1). However the periodization method requires that the integration is done over the interval $[0, 1]$ and that the singularity is located at one extremum of this

Table 1. Benchmark of the periodization method for a logarithmic singularity (from Helluy et al. (1998)).

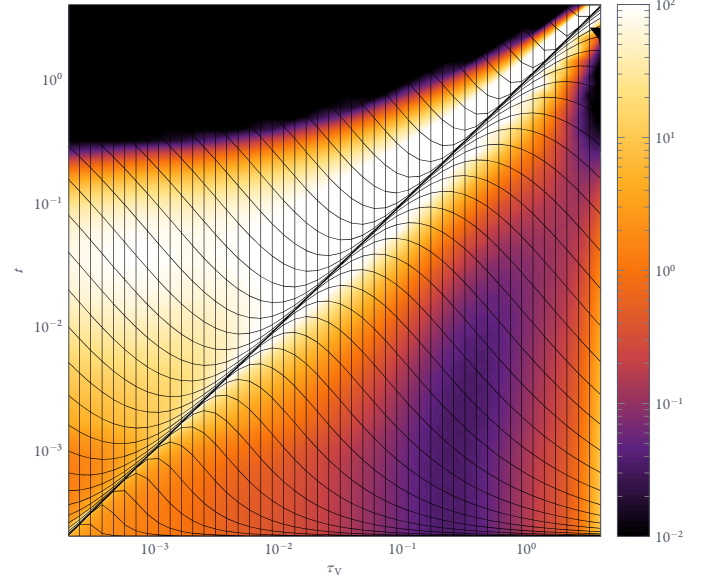| $f(x)$	k	C_k	γ	$|E_{10}(f)|$	$|E_{20}(f)|$
$\ln(x)$	2	3.22×10^0	3.03	0.30×10^{-2}	0.37×10^{-5}
	3	3.34×10^2	5.05	0.31×10^{-3}	0.93×10^{-5}
	5	3.05×10^4	9.09	0.14×10^{-4}	0.42×10^{-7}
	10	1.74×10^{14}	19.1	0.48×10^{-5}	0.92×10^{-11}

Fig. 3. Map of the integrand of $\bar{\mathcal{J}}_{ij}$ for a radiative transition (see Eq. 12c). The diagonal is singular, and the quadrature in the periodization method produces a mesh refinement around it. The color encoding of the integrant is in logarithmic arbitrary units.

interval. We thus split the integration domain into two subdomains, $[\tau_V^{in}, \tau_V]$ and $[\tau_V, \tau_V^{out}]$, and then normalize each subdomain to $[0, 1]$. However, since the MPI strategy limits the number of model layers handled by each process, this can lead to the undersampling of one of the subdomains, when τ_V approaches one of the bounds of the total domain (i.e., the boundaries of the model atmosphere). We have thus tested this method under the same conditions as those met in our problem, without any resampling of the subdomains. The result is displayed in Fig. 4. As expected, the relative error strongly increases when one of the subdomains includes too few model layers (typically fewer than 4) and becomes unacceptable for the extreme points. They will thus be considered ghost points in all subsequent computations. This does not affect the computation of the statistical equilibrium. Because we can compute the nonlinear function $\mathbf{F}(\mathbf{X})$ with good accuracy, the next step to find its root is to compute the Jacobian matrix to solve the system through a Newton's scheme.

3.3. Computation of the Jacobian matrix

The method presented here, and more generally the solving of nonlinear systems of equations, requires the computation of the Jacobian matrix. One of the advantages of the method we develop is that an analytical form can be derived. Here, the function $\mathbf{F}(\mathbf{X})$ is a set of $N \times P$ equations, N being the number of energy levels of the species considered, and P the number of layers in
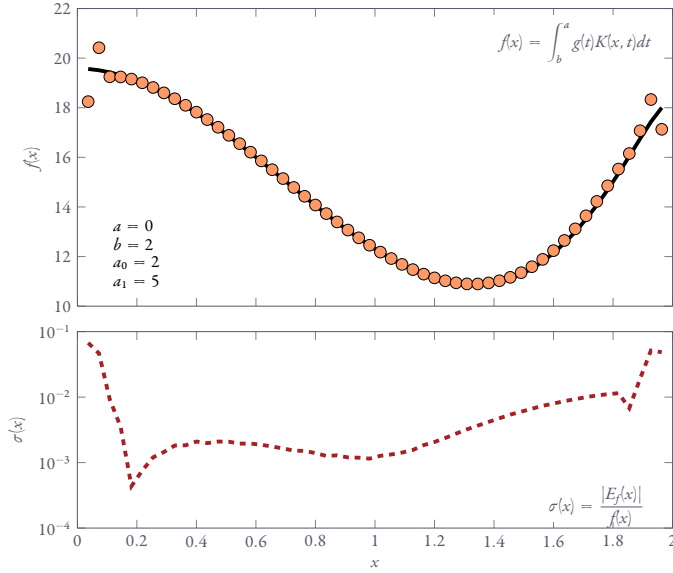
Fig. 4. Benchmarking of the periodization method applied to the integration with a logarithmically singular kernel ($g(t) = a_0 + a_1 t$, $K(x, t) = -\gamma - ln(|G(t) - G(x)|)$ and $G(x) = \int_0^t g(s)ds$). The integration domain has been split into two subdomains (see text for details).

the discretized model atmosphere. The differentiation of F_i^l over x_k^m (i and k referring to energy levels, and l and m to atmosphere layers) leads to a set of integro-differential equations, with the following four cases:

if $k = i$ and $l = m$ (31a)

$$\frac{\partial F_i^l}{\partial x_i^l} = \sum_{j \neq i}^{N} \frac{\partial \bar{\mathcal{J}}_{ij}^l}{\partial x_i^l} \left(x_i^l B_{ij} - x_j^l B_{ji} \right) + \sum_{j \neq i}^{N} \left(A_{ji} + B_{ji} \bar{\mathcal{J}}_{ji} + n_{col} C_{ji} \right),$$

if $k \neq i$ and $l = m$ (31b)

$$\frac{\partial F_i^l}{\partial x_k^l} = \frac{\partial \bar{\mathcal{J}}_{ik}^l}{\partial x_k^l} \left(x_k B_{ki} - x_i B_{ik} \right) - \left(A_{ki} + B_{ki} \bar{\mathcal{J}}_{ki} + n_{col} C_{ki} \right),$$

if $k = i$ and $l \neq m$ (31c)

$$\frac{\partial F_i^l}{\partial x_i^m} = \sum_{j \neq i}^{N} \frac{\partial \bar{\mathcal{J}}_{ij}^l}{\partial x_i^m} \left(x_i B_{ij} - x_j B_{ji} \right),$$

if $k \neq i$ and $t \neq m$ (31d)

$$\frac{\partial F_i^l}{\partial x_k^m} = \frac{\partial \bar{\mathcal{J}}_{ik}^l}{\partial x_k^m} \left(x_k B_{ki} - x_i B_{ik} \right).$$

Equations (31a) and (31b) represent local couplings. Mathematically, they correspond to the diagonal and block-diagonal parts of the Jacobian, respectively. Physically, Eq. (31a) describes perturbations of the statistical equilibrium of a given level, by modifying the population of this level, within one atmospheric layer. Equation (31b) describes the perturbations of the statistical equilibrium of a given level, by modifying the population of other levels, still within one atmospheric layer. Equations (31c) and (31d) correspond to off-diagonal blocks and describe nonlocal perturbations, i.e., the influence of the populations in another layer. Naturally, in these last differentiations the coupling is purely radiative, whereas the first and second differentiations include collisional processes.

To determine the terms $\partial \bar{\mathcal{J}}_{ik}^l / \partial x_k^m$, one needs to differentiate an integral expression over the variable x_k^m. This derivation must

be considered a functional derivative, and be performed before discretization (see Appendix B for the demonstration),

$$\frac{\delta \bar{\mathcal{J}}_{ij} [x_k] (\tau_V)}{\delta x_k (s)} = \lim_{\varepsilon \to 0} \frac{\bar{\mathcal{J}}_{ij} [x_k(\tau_V) + \varepsilon \delta(\tau_V - s)] - \bar{\mathcal{J}}_{ij} [x_k(\tau_V)]}{\varepsilon},$$
(32)

where $\delta(\tau_V - s)$ is the Dirac distrbution.

In the case $\tau_V \neq s$ ($l \neq m$), the derivatives for each component are

$$\frac{\delta \bar{\mathcal{J}}_{ij}^{\text{Line}}[x_i, x_j](\tau_V)}{\delta x_i(s)} = D_{ij}\zeta(s)K(s, \tau_V) + D_{ij}\zeta(s)E_{ij}$$

$$\times \begin{cases} \int_k^{\tau_V^{\text{Out}}} \zeta(t)x_i(t)\tilde{K}_0(t, \tau_V, s)dt & \text{if } k > \tau_V \\ \int_{\tau_V^{\text{In}}}^k \zeta(t)x_i(t)\tilde{K}_0(t, \tau_V, s)dt & \text{if } k < \tau_V \end{cases},$$
(33a)

$$\frac{\delta \bar{\mathcal{J}}_{ij}^{\text{Line}}[x_i, x_j](\tau_V)}{\delta x_j(s)} = -D_{ij}\zeta(s)E_{ij}$$

$$\times \begin{cases} \int_k^{\tau_V^{\text{Out}}} \zeta(t)x_i(t)\tilde{K}_0(t, \tau_V, s)dt & \text{if } k > \tau_V \\ \int_{\tau_V^{\text{In}}}^k \zeta(t)x_i(t)\tilde{K}_0(t, \tau_V, s)dt & \text{if } k < \tau_V \end{cases},$$
(33b)

$$\frac{\delta \bar{\mathcal{J}}_{ij}^{\text{Cont}}[x_i, x_j](\tau_V)}{\delta x_i(s)} = \zeta(s)E_{ij}$$

$$\times \begin{cases} \int_k^{\tau_V^{\text{Out}}} \xi_{ij}(t)B_{ij}(t)\tilde{L}_0(t, \tau_V, s)dt & \text{if } k > \tau_V \\ \int_{\tau_V^{\text{In}}}^k \xi_{ij}(t)B_{ij}(t)\tilde{L}_0(t, \tau_V, s)dt & \text{if } k < \tau_V \end{cases},$$
(33c)

$$\frac{\delta \bar{\mathcal{J}}_{ij}^{\text{Cont}}[x_i, x_j](\tau_V)}{\delta x_j(s)} = -\zeta(s)E_{ij}$$

$$\times \begin{cases} \int_k^{\tau_V^{\text{Out}}} \xi_{ij}(t)B_{ij}(t)\tilde{L}_0(t, \tau_V, s)dt & \text{if } k > \tau_V \\ \int_{\tau_V^{\text{In}}}^k \xi_{ij}(t)B_{ij}(t)\tilde{L}_0(t, \tau_V, s)dt & \text{if } k < \tau_V \end{cases}.$$
(33d)

However $\bar{\mathcal{J}}$ is not differentiable at $x(t) = x(\tau_V)$ because of the singularity, therefore Eq. (33a) is undetermined for $s = \tau_V$. For these local terms, we thus need an approximate expression of the radiation field. We emphasize that we only apply this approximation to the computation of the Jacobian matrix.

3.4. Sparsity of Jacobian matrix

In Section 3.1, we argue that GMRES is a well-adapted method for large-scale problems if the Jacobian matrix is sparse, as it only requires matrix-vector products, and no matrix inversion. The sparsity (or density) coefficient is defined by the ratio of nonzero elements over the total number of elements. The Jacobian matrix is composed by $P \times P$ submatrix of $N \times N$ elements leading to $NP \times NP$ elements. The block diagonal of this matrix
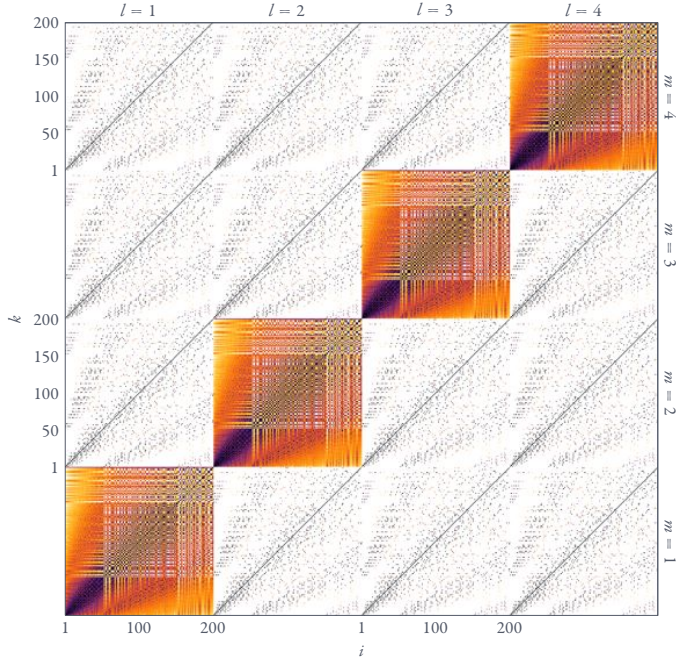
Fig. 5. Pattern of the Jacobian matrix for 4 model atmosphere layers and 200 energy levels.

is full because of the irreducibility of the collisional coupling within a given layer (no selection rules for collisions). The number of nonzero elements in the block-diagonal part is $P \times N^2$. In the $P(P-1)$ off-diagonal submatrices, the nonzero elements are only due to radiative coupling between different layers, and their number is thus linked to the selection rules for allowed transitions. These rules thus lead to a sparsity of each submatrix with a density $\rho_{\mathrm{mol}} = N_{\mathrm{transitions}}/N^2$. The density of the complete Jacobian matrix is then given by

$$\rho = \frac{P \times N^2 + P(P-1) \times N_{\mathrm{transitions}}}{N^2 P^2} = \frac{1 + \rho_{\mathrm{mol}}(P-1)}{P}. \quad (34)$$

Then, as $\rho_{\mathrm{mol}}(N)$ decreases rapidly with N (e.g., $\rho_{H_2O} \sim 10^{-3}$,
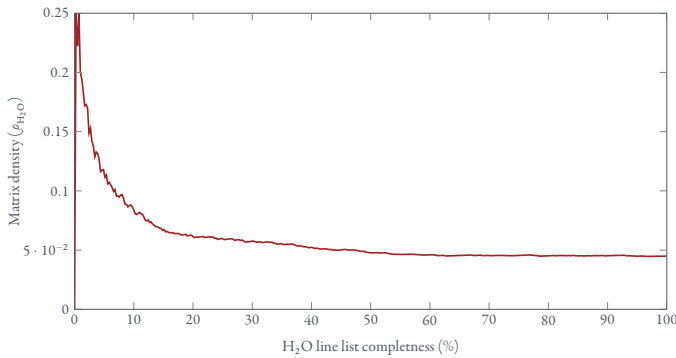

Fig. 6. Example of the evolution of the off-diagonal, block-matrix sparsity (ρ_{mol}) according to the number of considered levels. This example is computed for the 411 levels of the ortho–H_2O molecule.

for $N = 411$, see Figs. 6 and 7) and ρ is inversely proportional to P, the larger the system, the sparser the Jacobian matrix. The GMRES scheme is thus well adapted and asymptotically approaches the NP complexity.
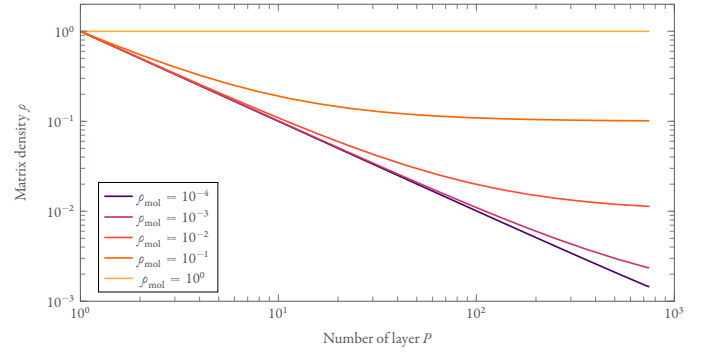
Fig. 7. Evolution of the Jacobian matrix sparsity according to the number of atmospheric layers for different values of ρ_{mol}.

3.5. Jacobian block-diagonal approximation

Assuming that $\tau_T(\tau_V) - \tau_T(t) \propto (t - \tau_V)$ (first order linearization) when t is close to τ_V, because the kernel decreases when $|t - \tau_V|$ increases, we assume that every function at t is equal to its value at τ_V.

Splitting the integral over t at τ_V and inverting the integrals over t and μ, Gonzalez Garcia et al. (2008) show that the mean intensity can be rewritten as

$$\bar{\mathcal{J}}_{ij}(\tau_V) = \frac{x_i(\tau_V)}{x_j(\tau_V) - x_i(\tau_V)} (1 - \beta_-(\tau_V) - \beta_+(\tau_V)) \quad (35a)$$

$$\bar{\mathcal{J}}_{ij}(\tau_V) = \tau_{max} I_\nu \beta(\tau_V), \quad (35b)$$

where $\beta_-(\tau_V)$ and $\beta_+(\tau_V)$ are inward and outward escape probabilities (defined in Gonzalez Garcia et al. (2008)). The $\beta(\tau_V)$ invoke the function $E_2(x)$, which is not singular and differentiable everywhere in the domain. This enables an analytic estimate of the bloc diagonal terms. The pattern of the Jacobian is displayed in Fig. 5.

3.6. The numerical code MOrad

The above strategy has been implemented in a numerical code named MOrad written in Fortran 90. It takes as inputs:

1. A model atmosphere (temperature, atomic, molecular, and electronic densities etc. as a function of geometrical thickness)
2. Spectroscopic data (level energies, statistical weights, radiative rates)
3. Collisional rates

Then it computes the continuous opacities and uses an initial guess to calculate the function (22) and its Jacobian matrix. Choosing the initial guess $\mathbf{X}_0$ is an important operation. In MOrad we tested the use of an LTE solution (Boltzmann distribution), and a 0 K solution (the molecules in the ground level). We find that the convergence rate is better with the "0 K" initial guess.

The singular integrations needed to compute $\bar{\mathcal{J}}$ are performed with the periodization method. The Green's function needed to compute $\bar{\mathcal{J}}$ is stored at each point. The required space scales as the number of layers times the number of frequency points (in a parallel code; otherwise it would scale with the square of the number of layers). The evaluation of the function and the Jacobian are parallelized with MPI. At first glance, two parallelization strategies may be considered: either a parallelization over frequencies or over atmosphere layers. However, van

Noort et al. (2002) show that the most efficient parallelization to reduce communications is spatial parallelization because of the strong local coupling between energy levels. The populations of all energy levels within each atmosphere layer should thus be computed by the same processor, with several adjacent layers possibly treated by the same processor. In MOrad we adopt a subdomain decomposition where each processor computes the physical values needed in one given layer and solves the local statistical equilibrium. Integrations are performed with a scalable algorithm, which minimizes global communications to preserve a good scalability.

The code has been interfaced with a nonlinear solver of the parallel and scalable library PETSc (see, e.g., Balay et al. 2013b,a, 1997). This solver takes as input the distributed function vector and the Jacobian matrix and returns the root of the function after preconditioning.

The main advantages of this library are its performance and its scalability. Moreover, PETSc provides access to many different nonlinear methods and preconditioning, which allows us a versatile choice of the best method to be adapted for a given physical problem. Other efficient libraries, such as the parallel IO library HDF5, are interfaced with MOrad to preserve the performances for large-scale problems and the portability of output data. Future planned developments include an MPI/OpenMP hybridization (openMP parallelization over frequencies).

4. Discussion

4.1. Example of application

To illustrate the applicability of MOrad, we compute the nonLTE departure coefficients of water molecules in a MARCS (Gustafsson et al. 2008) red supergiant model atmosphere. The model considered here has an effective temperature of 3500 K, a $\log g = 0$, a solar metallicity, and a microturbulence of 2 km s^{-1}. These parameters correspond to typical red supergiant stars. We consider more than 800 rovibrational levels, leading to more than 330 000 transitions and 15 000 lines (see the Grotrian diagram of ortho H_2O in Fig. 8). The energy levels and the radiative coefficients are taken from Barber et al. (2006). The collisional rates for H_2O–H_2 and H_2O–e^- are taken from Faure & Josselin (2008).
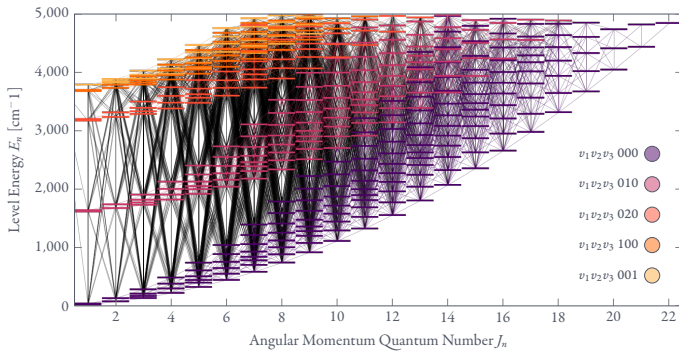
Fig. 8. Grotrian diagram of the ortho–H_2O molecule showing the radiative transitions taken into account. The parameter J_n is the rotational angular momentum of H_2O. Colors indicate different vibrational states.

The code MOrad was launched on 47 processors using a GMRES Newton-Krylov method and a line-search global convergence strategy. We conducted preconditioning with the AMS method (Block Additive Schwarz method), where each subblock is preconditioned with an approximate iterative LU factorization

method. We obtain the root $\mathbf{X}^\star$ of the function ($\|\mathbf{F}(\mathbf{X}^\star)\|_\infty \leq 10^{-12}$) after three nonlinear iterations and a total time of code execution of less than ~1h on the Alarik LUNARC system (see Lunarc 2013). The departure coefficients are presented in Fig. 9. The convergence rate is Q-quadratic, though a quadratic convergence rate would be expected. This may be because of an over reconditioning or a problem that is too simple(close to linear), but we obtain the same Q-quadratic convergence for a purely radiative problem.
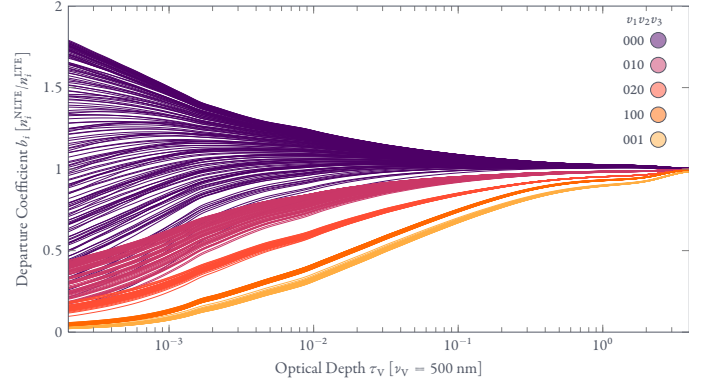
Fig. 9. LTE departure coefficients of ortho-H_2O in a red supergiant model atmosphere.

A detailed analysis of the results is out of the scope of this paper and is devoted to a forthcoming paper; a preliminary analysis was done in Lambert et al. (2013). In particular a detailed discussion of the use of the super-level approximation, which seems appropriate for the vibrationally excited states, will be presented. In short, the non-LTE calculations lead to stronger lines, especially rotational lines in the fundamental vibrational state around 2 μm, compared to LTE computation.

4.2. Performance and complementarity of the method

The method presented here is naturally accurate as it is equivalent to finding the root of a function and thus avoids any false convergence (assuming the solution is unique)[3]. Furthermore, we emphasize that because the integration is analytic, the method leads naturally to null residue. Concerning the performance of the method, the performances of the method are difficult to assess. The convergence rate depends upon the choice of both the solver and the preconditioning. These choices are not unique and may vary according to the considered problem (species and model atmosphere). The memory demands scales with the dimensionality of the considered problem.

The computation of the function and the Jacobian matrix leads to a scheme with an asymptotic complexity in $O(Nb_{\mathrm{freq}} \times (N \times P)^2)$. The complexity of the resolution scheme is more difficult to evaluate because the preconditioned Newton-Krylov method has a cost in $N_{\mathrm{freq}} \times N^2 \times P^2$ (N_{freq} being the number of frequency points sampling the line profile, N the number of energy levels, and P the number of layers in the discretized atmosphere), but invokes another term that evolves as the square

---

[3] The solution of the Eddington problem (the two-level atom in the isothermal optically thick medium) is recovered at the 1st iteration with preconditioning, and after five iterations without preconditioning (relative error < 0.005 for $\epsilon = 10^{-5}$ and $\tau_{max} = 1000$). This is expected because the inversion of the matrix by the Krylov subspace method is well known to work in this kind of situation.

of the iteration index. This term is partially eliminated by restarting GMRES and is not really important in practice because the number of iterations is small after appropriate preconditioning.

Furthermore, the performance is better than $O(n^2)$ and probably close to $O(n)$. Indeed, for a sparse matrix, GMRES evolves in $O(n)$. Because the only nonzero elements of the off-diagonal blocks of the Jacobian matrix are due to radiative coupling between layers, the sparsity of J evolves as the ratio of the number of radiative transitions over N^2. Taking a complexity in $Nb_{\text{freq}} \times N^2 \times P^2$ is then an superior limit, which underevaluates the method scalability in real cases.

As a comparison, the computation of the equivalent function of $\mathbf{F}(\mathbf{X})$ in MULTI has a complexity evolving in $N \times P$ because of the Scharmer operator, which performs no numerical integration to evaluate diffusion integrals but only one quadrature point. The system is solved with a Newton-Raphson scheme, which has a cost in $N^3 \times P^2$ when using the full matrix or $N^3 \times P$ when using the local operator (block-diagonal matrix; M. Carlsson, priv. comm.). The theoretical asymptotic speed-up for MOrad compared to MULTI are summarized in Table 4.2.

| | $t_{\text{MULTI}}/t_{\text{MOrad}}$ | |
	Full Operator	OAB (diagonal)
$\rho = 1$	N	N/P
$\rho = 0$	PN^2	N^2

Table 2. Theoretical asymptotic speed-up for two extreme cases in terms of $\rho = \text{sparsity}(\mathbf{J})(0 \le \rho \le 1)$ for MOrad, compared to MULTI with either full or local operator (OAB stands for Olson, Auer, & Buchler Operator).

From Table 4.2, one can see that our method is complementary to existing methods, such as direct linearization. Indeed, in some cases, with a large number of levels and a small number of layers, the expected time of computation is shorter.

5. Conclusion

We have developed a new method to solve the RTE in non-LTE conditions. Rather than solving the RTE itself, we solve the statistical equilibrium equation system, which includes an analytical formal solution of the RTE. One of the main advantages of this choice is that the coefficients of the associated Jacobian matrix can be computed exactly. With this approach, a nonlinear solver can be used. In other words, this method avoids any linearization or stationary iterative schemes, which may converge very slowly or even falsely in some extreme cases (a very low non-LTE parameter ϵ and/or highly discretized media). We chose a method based on the GMRES nonlinear method. It has been implemented in a code that uses the PETSc library. We pay special attention to the proper treatment of the singularity that arises in the integration of the mean radiation field.

Furthermore, this method is parallelized using MPI, which makes it applicable to large-scale systems such as molecules and atoms in stellar atmospheres or interstellar medium, as it has an excellent scalability with the number of energy levels. Indeed, our code has been successfully applied to the modeling of water in the atmospheres of red supergiants. The results will be presented in a forthcoming paper. Further developments of the code include hybrid parallelization, with openMP parallelization over frequencies. We then plan to make the code public.

References

Ahues, M., D'Almeida, F., Largillier, A., Titaud, O., & Vasconcelos, P. 2002, Journal of Computational and Applied Mathematics, 140, 13

Auer, L. 1991, in NATO ASIC Proc. 341: Stellar Atmospheres - Beyond Classical Models, ed. L. Crivellari, I. Hubeny, & D. G. Hummer, 9

Balay, S., Brown, J., , et al. 2013a, PETSc Users Manual, Tech. Rep. ANL-95/11 - Revision 3.4, Argonne National Laboratory

Balay, S., Brown, J., Buschelman, K., et al. 2013b, PETSc Web page, http://www.mcs.anl.gov/petsc

Balay, S., Gropp, W. D., McInnes, L. C., & Smith, B. F. 1997, in Modern Software Tools in Scientific Computing, ed. E. Arge, A. M. Bruaset, & H. P. Langtangen (Birkhäuser Press), 163–202

Barber, R. J., Tennyson, J., Harris, G. J., & Tolchenov, R. N. 2006, MNRAS, 368, 1087

Cannon, C. J. 1973, J. Quant. Spec. Radiat. Transf., 13, 1011

Castor, J. I. 1970, MNRAS, 149, 111

Damgaard, P. H., Hjorth, P. G., & Thejll, P. A. 1992, A&A, 254, 422

Dickel, H. R. & Auer, L. H. 1994, ApJ, 437, 222

Fabiani Bendicho, P., Trujillo Bueno, J., & Auer, L. 1997, A&A, 324, 161

Faure, A. & Josselin, E. 2008, A&A, 492, 257

Gonzalez Garcia, M., Le Bourlot, J., Le Petit, F., & Roueff, E. 2008, A&A, 485, 127

Gustafsson, B., Edvardsson, B., Eriksson, K., et al. 2008, A&A, 486, 951

Hauschildt, P. H., Starrfield, S., Shore, S. N., Allard, F., & Baron, E. 1995, ApJ, 447, 829

Helluy, P., Maire, S., & Ravel, P. 1998, Academie des Sciences Paris Comptes Rendus Serie Sciences Mathematiques, 327, 843

Hubeny, I. & Mihalas, D. 2014, Theory of Stellar Atmospheres

Juvela, M. 2005, A&A, 440, 531

Klein, R. I., Castor, J. I., Dykema, P. G., Greenbaum, A., & Taylor, D. 1989, J. Quant. Spec. Radiat. Transf., 41, 199

Lambert, J., Josselin, E., Ryde, N., & Faure, A. 2013, in EAS Publications Series, Vol. 60, EAS Publications Series, ed. P. Kervella, T. Le Bertre, & G. Perrin, 111–119

Lunarc. 2013, Alarik system details, http://www.lunarc.lu.se/Systems/AlarikDetails

Ng, K.-C. 1974, J. Chem. Phys., 61, 2680

Olson, G. L., Auer, L. H., & Buchler, J. R. 1986, J. Quant. Spec. Radiat. Transf., 35, 431

Paletou, F. & Anterrieu, E. 2009, A&A, 507, 1815

Rybicki, G. B. 1997, ApJ, 479, 357

Scharmer, G. B. & Carlsson, M. 1985, Journal of Computational Physics, 59, 56

Socas-Navarro, H. & Trujillo Bueno, J. 1997, ApJ, 490, 383

Trujillo Bueno, J. & Fabiani Bendicho, P. 1995, ApJ, 455, 646

Štěpán, J. & Trujillo Bueno, J. 2013, A&A, 557, A143

van Noort, M., Hubeny, I., & Lanz, T. 2002, ApJ, 568, 1066

Appendix A: Formulation

One starts with the classical formulation of the radiative transfer equation in a plane parallel geometry,

$$\mu \frac{dI_\nu}{ds} = -\left(\chi_\nu^{\text{Cont}} + \chi_\nu^{\text{Line}}\right) I_\nu + \eta_\nu^{\text{Cont}} + \eta_\nu^{\text{Line}}. \tag{A.1}$$

The monochromatic extinction and emissivity coefficients are expressed with an explicit dependence on levels populations as follows:

$$\chi_\nu^{\text{Line}} = \phi_\nu^{ul} \chi_{ul} = \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}(B_{lu}n_l - B_{ul}n_u) \tag{A.2a}$$

$$\eta_\nu^{\text{Line}} = \phi_\nu^{ul} \eta_{ul} = \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}A_{ul}n_u. \tag{A.2b}$$

By inserting relative populations $f_i = n_i/n^k$, where n^k is the total density of the species labeled by k, one gets

$$\chi_\nu^{\text{Line}} = \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}n^k(B_{lu}f_l - B_{ul}f_u) \tag{A.3a}$$

$$\eta_\nu^{\text{Line}} = \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}n^k A_{ul}f_u. \tag{A.3b}$$

One brings up the new variable $x_i = f_i/g_i$, and using the relation $B_{ul}g_u = B_{lu}g_l$, where g_i is the statistical weight of the level i, yields

$$\chi_\nu^{\text{Line}} = \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}n^k g_u (B_{lu}\frac{f_l}{g_u} - B_{ul}\frac{f_u}{g_u}) = \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}n^k g_u B_{ul}(x_l - x_u) \tag{A.4a}$$

$$\eta_\nu^{\text{Line}} = \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}n^k g_u A_{ul}\frac{f_u}{g_u} = \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}n^k g_u A_{ul}x_u. \tag{A.4b}$$

Including the formulation of χ_ν^{Line} and η_ν^{Line} in the radiative transfer equation, one obtains

$$\mu \frac{dI_\nu}{ds} = -\left(\chi_\nu^{\text{Cont}} + \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}n^k g_u B_{ul}(x_l - x_u)\right) I_\nu + \eta_\nu^{\text{Cont}} + \phi_\nu^{ul} \frac{h\nu_{ul}}{4\pi}n^k g_u A_{ul}x_u. \tag{A.5}$$

In order to create a transition independent scale, the previous equation is divided by an arbitrarily chosen continuum extinction coefficient χ_V, here $\chi_V \equiv \chi_{\nu=500nm}$. Moreover we express it per hydrogen atom to bring out the ratio n^k/n^{H} for numerical reasons. This normalization is symbolized by a tilde ($\sim$) for other variables,

$$\mu \frac{dI_\nu}{\chi_V ds} = -\left(\frac{\tilde{\chi}_\nu^{\text{Cont}}}{\tilde{\chi}_V}\frac{n^{\text{H}}}{n^{\text{H}}} + \frac{h\nu_{ul}}{4\pi}\frac{n^k}{n^{\text{H}}}\frac{1}{\tilde{\chi}_V}B_{ul}g_u(x_l - x_u)\phi_\nu^{ul}\right) I_\nu + \frac{h\nu_{ul}}{4\pi}\frac{n^k}{n^{\text{H}}}\frac{1}{\tilde{\chi}_V}A_{ul}g_u x_u \phi_\nu^{ul} + \frac{\eta_\nu^{\text{Cont}}}{n^{\text{H}}\tilde{\chi}_V}\frac{n^{\text{H}}\tilde{\chi}_\nu^{\text{Cont}}}{\chi_\nu^{\text{Cont}}}. \tag{A.6}$$

With $\xi_\nu = \frac{\tilde{\chi}_\nu}{\tilde{\chi}_V}$, $E_{ij} = \frac{h\nu_{ij}}{4\pi}B_{ij}g_i$, $\zeta = \frac{n^k}{n^{\text{H}}}\frac{1}{\tilde{\chi}_V}$, $D_{ij} = \frac{h\nu_{ij}}{4\pi}A_{ij}g_i$, $S_\nu = \frac{\eta_\nu^{\text{Cont}}}{\chi_\nu^{\text{Cont}}}$, and $d\tau_V = -\chi_V ds$. Assuming that the continuum is formed in LTE ($S_\nu^{\text{Cont}} = B_\nu$), one obtains

$$\mu \frac{dI_\nu}{d\tau_V} = \left[\xi_\nu + E_{ul}\zeta\phi_\nu^{ul}(x_l - x_u)\right] I_\nu - D_{ul}x_u\zeta\phi_\nu^{ul} - \xi_\nu B_\nu. \tag{A.7}$$

Appendix B: Jacobian matrix coefficients

We start with the analytic formulation of the mean intensity field for the line component (similar demonstration for contunium), i.e.,

$$\bar{\mathcal{J}}_{ij}[x_i, x_j](\tau_V) = D_{ij} \int_{\tau_V^{\text{In}}}^{\tau_V^{\text{Out}}} \zeta(t)x_i(t)K[x_i, x_j](t, \tau_V)dt. \tag{B.1}$$

Because the functional derivative of a functional product is

$$\frac{\delta(F[y]G[y])}{\delta y(x)} = G[y]\frac{\delta F[y]}{\delta y(x)} + F[y]\frac{\delta G[y]}{\delta y(x)}, \tag{B.2}$$

we functionally derive the product $x_i(t)K[x_i, x_j]$ over $x_i(\tau_V = s)$, which yields,

$$\frac{\delta\bar{\mathcal{J}}_{ij}[x_i, x_j](\tau_V)}{\delta x_i(s)} = D_{ij} \int_{\tau_V^{\text{In}}}^{\tau_V^{\text{Out}}} \zeta(t)\frac{\delta x_i(t)}{\delta x_i(s)}K[x_i, x_j](t, \tau_V)dt + D_{ij} \int_{\tau_V^{\text{In}}}^{\tau_V^{\text{Out}}} \zeta(t)x_i(t)\frac{\delta K[x_i, x_j](t, \tau_V)}{\delta x_i(s)}dt. \tag{B.3}$$

Equation (B.3) can be rewritten as a sum of two integrals over t, I_1 and I_2.
The first term I_1 simplifies

$$I_1 = D_{ij} \int_{\tau_V^{In}}^{\tau_V^{Out}} \zeta(t) \frac{\delta x_i(t)}{\delta x_i(s)} K[x_i, x_j](t, \tau_V) dt = D_{ij} \int_{\tau_V^{In}}^{\tau_V^{Out}} \zeta(t) \delta(t - s) K[x_i, x_j](t, \tau_V) dt$$
$$= D_{ij} \zeta(s) K(s, \tau_V) \qquad . \tag{B.4}$$

This term is an indefinite form if $s = \tau_V$, because the kernel $K(x, y)$ is singular at $x = y$.
We consider here the case $s \neq \tau_V$ (the case $s = \tau_V$ is detailed in Section 3.4).
The second term I_2 requires the computation of the functional derivative of the kernel, which is

$$K[x_i, x_j](t, \tau_V) = \frac{1}{2} \int_0^\infty \phi_v(\tau_V) \phi_v(t) E_1 \left(\Delta \tau_v^{Tot}[x_i, x_j](t, \tau_V) \right) dv. \tag{B.5}$$

With the chain rule theorem of functional derivatives, we get

$$\frac{\delta f(F[y])}{\delta y(x)} = \frac{df(F[y])}{dF[y]} \frac{\delta F[y]}{\delta y(x)}. \tag{B.6}$$

We then derive the kernel K, finding

$$\frac{\delta K[x_i, x_j](t, \tau_V)}{\delta x_i(s)} = \frac{1}{2} \int_0^\infty \phi_v(\tau_V) \phi_v(t) \times \frac{\delta E_1 \left(\Delta \tau_v^{Tot}[x_i, x_j](t, \tau_V) \right)}{\delta x_i(s)} dv$$
$$= -\frac{1}{2} \int_0^\infty \phi_v(\tau_V) \phi_v(t) \times E_0 \left(\Delta \tau_v^{Tot}[x_i, x_j](t, \tau_V) \right) \frac{\delta \Delta \tau_v^{Tot}[x_i, x_j](t, \tau_V)}{\delta x_i(s)} dv. \tag{B.7}$$

To simplify the calculus, we rewrite $\Delta \tau_v^{Tot}$ as

$$\Delta \tau_v^{Tot}[x_i, x_j](t, \tau_V) = \left| \int_t^{\tau_V} \xi_v(s') + E_{ij} \zeta(s') \phi_v(s') \left[x_j(s') - x_i(s') \right] ds' \right|$$
$$= \int_{\tau_V^{In}}^{\tau_V^{Out}} \Theta(t, \tau_V, s') \left[\xi_v(s') + E_{ij} \zeta(s') \phi_v(s') \left[x_j(s') - x_i(s') \right] \right] ds', \tag{B.8}$$

where the distribution $\Theta(t, \tau_V, s) = -2(\theta(t - \tau_V) - 0.5)(\theta(s - t) - \theta(s - \tau_V))$, and $\theta(x)$ is the Heaviside function. This distribution can be rewritten,

$$\Theta(t, \tau_V, s) = \begin{cases} \Pi_{a,s}(t) & \text{if } s < \tau_V \\ \Pi_{s,b}(t) & \text{if } s > \tau_V \end{cases}. \tag{B.9}$$

We functionally derive Eq. (B.8), yielding

$$\frac{\delta \Delta \tau_v^{Tot}[x_i, x_j](t, \tau_V)}{\delta x_i(s)} = -\int_{\tau_V^{In}}^{\tau_V^{Out}} \Theta(t, \tau_V, s') E_{ij} \zeta(s') \phi_v(s') \frac{\delta x_i(s')}{\delta x_i(s)} ds' = -\int_{\tau_V^{In}}^{\tau_V^{Out}} \Theta(t, \tau_V, s') E_{ij} \zeta(s') \phi_v(s') \delta(s - s') ds'$$
$$= -E_{ij} \zeta(s) \Theta(t, \tau_V, s) \phi_v(s). \tag{B.10}$$

We insert this expression in Eq. (B.7) and we define the modified kernel $\tilde{K}_0(t, \tau_V, s)$,

$$\frac{\delta K[x_i, x_j](t, \tau_V)}{\delta x_i(s)} = E_{ij} \zeta(s) \Theta(t, \tau_V, s) \frac{1}{2} \int_0^\infty \phi_v(\tau_V) \phi_v(t) \phi_v(s) E_0 \left(\Delta \tau_v^{Tot}[x_i, x_j](t, \tau_V) \right) dv$$
$$= E_{ij} \zeta(s) \Theta(t, \tau_V, s) \tilde{K}_0[x_i, x_j](t, \tau_V, s). \tag{B.11}$$

The integral I_2 from equation (B.3) can be rewritten as

$$I_2 = D_{ij} \int_{\tau_V^{In}}^{\tau_V^{Out}} \zeta(t) x_i(t) \frac{\delta K[x_i, x_j](t, \tau_V)}{\delta x_i(s)} dt$$
$$= D_{ij} E_{ij} \zeta(s) \int_{\tau_V^{In}}^{\tau_V^{Out}} \Theta(t, \tau_V, s) \zeta(t) x_i(t) \tilde{K}_0[x_i, x_j](t, \tau_V, s) dt, \tag{B.12}$$

which, according to the Θ properties, becomes

$$I_2 = D_{ij} E_{ij} \zeta(s) \times \begin{cases} \int_s^{\tau_V^{Out}} \zeta(t) x_i(t) \tilde{K}_0(t, \tau_V, s) dt & \text{if } s > \tau_V \\ \int_{\tau_V^{In}}^s \zeta(t) x_i(t) \tilde{K}_0(t, \tau_V, s) dt & \text{if } s < \tau_V \end{cases}. \tag{B.13}$$

Finally, we get the derivative of the mean radiation field,

$$\frac{\delta \bar{\mathcal{J}}_{ij}[x_i, x_j](\tau_V)}{\delta x_i(s)} = D_{ij}\zeta(s)\left(K(s,\tau_V) + E_{ij} \times \begin{cases} \int_{s,s}^{\tau_V^{\text{Out}}} \zeta(t)x_i(t)\tilde{K}_0(t,\tau_V,s)dt & \text{if } s > \tau_V \\ \int_{\tau_V^{\text{In}}} \zeta(t)x_i(t)\tilde{K}_0(t,\tau_V,s)dt & \text{if } s < \tau_V \end{cases}\right). \tag{B.14}$$

Appendix C: Definition of the variables

Table C.1. Physical quantities

Symbol	Meaning	units (cgs)
μ	Direction cosine ($\mu = \cos(\theta)$)	
T	Temperature	K
n^k	Total density of species k	cm^{-3}
n_i	Density in an energy level labeled by i	cm^{-3}
f_i	Relative density ($f_i = n_i/n^k$)	
x_i	Weighted relative density ($x_i = f_i/g_i$)	
g_i	Statistical weight of level i	
ν	Frequency	Hz
I_ν	Specific intensity	$\text{erg s}^{-1}\,\text{cm}^{-2}\,\text{sr}^{-1}\,\text{Hz}^{-1}$
$\mathcal{J}_\nu$	Mean intensity	$\text{erg s}^{-1}\,\text{cm}^{-2}\,\text{Hz}^{-1}$
τ_ν	Optical depth	
τ_V	Optical depth at 500 nm	
S_ν	Source function	$\text{erg s}^{-1}\,\text{cm}^{-2}\,\text{sr}^{-1}\,\text{Hz}^{-1}$
B_ν	Planck function	$\text{erg s}^{-1}\,\text{cm}^{-2}\,\text{sr}^{-1}\,\text{Hz}^{-1}$
χ_ν	Monochromatic extinction	cm^{-1}
η_ν	Monochromatic emissivity	$\text{erg s}^{-1}\,\text{cm}^{-3}\,\text{sr}^{-1}\,\text{Hz}^{-1}$
β_ν	Escape probability	
A_{ul}	Einstein coefficient for spontaneous de-excitation	s^{-1}
B_{ij}	Einstein coefficient for stimulated (de)excitation	$\text{erg}^{-1}\,\text{cm}^2\,\text{Hz}$
C_{ij}	Rate for collisional (de)excitation	$\text{cm}^3\,\text{s}^{-1}$
ϕ_ν^{ij}	Normalized line profile of the transition $i \to j$	Hz^{-1}

Table C.2. Functions, operators and distributions

Functions	Name	Meaning
$E_n(x)$	n^{th} Exponential integral	$E_n = \int_1^\infty \frac{e^{-xt}}{t^n}dt$
$B_\nu(T)$	Planck function	$B_\nu(T) = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/k_B T} - 1}$
$\delta(x)$	Dirac distribution	$\delta(x) = \begin{cases} +\infty & \text{if } x = 0 \\ 0 & \text{if } x \neq 0. \end{cases}$ $\int_{-\infty}^{+\infty} \delta(x)dx = 1$
$\Pi(x)$	Gate distribution	$\Pi(x) = H(x + 1/2) - H(x - 1/2)$
$\Pi_{a,b}(x)$	Generalized gate distribution (boxcar)	$\Pi_{a,b}(x) = \Pi((x - (b + a)/2(b - a)))$
$\theta(x)$	Heaviside distribution	$\theta(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 0. \end{cases}$
$\mathcal{G}$	Green's function	$\mathcal{L}\mathcal{G}(x,s) = \delta(x - s)$
$\bar{f}$	Integration of f weighted by a line profile	$\bar{f} = \int_{-\infty}^{+\infty} f_\nu\phi_\nu d\nu$